2020 CENSUS PROGRAM MEMORANDUM SERIES: 2019.13

Date: July 1, 2019

MEMORANDUM FOR: The Record

From: Albert E. Fontenot, Jr. **(Signed July 1, 2019)**
                             Associate Director for Decennial Census Programs

Subject: Disclosure Avoidance System Design Parameters and Global Privacy-Loss Budget
                             for the 2018 End-to-End Census Test

Contact: Cynthia Hollingsworth
                             Decennial Census Management Division
                             301-763-3655
                             cynthia.davis.hollingsworth@census.gov

                             Robert Sienkiewicz
                             Center for Enterprise Dissemination
                             301-763-1234
                             Robert.sienkiewicz@census.gov

This memorandum documents requirements the 2020 Census Program has received from the Data Stewardship Executive Policy Committee (DSEP) regarding how to protect the information we collect from the American public during the 2020 Census. DSEP has instructed us to apply differentially private disclosure avoidance methods to all of the statistics we produce from these data, starting with data collected during the 2018 End-to-End (2018 E2E) Census Test. DSEP will manage the amount of privacy loss associated with the 2020 Census data products, called the "privacy-loss budget," and make policy decisions regarding the design of the 2020 Disclosure Avoidance System (DAS). I will communicate these policy decisions and their associated requirements as we receive them from DSEP through the 2020 Census Memorandum Series.

The Census Bureau has continually added better and stronger protections to keep the data it publishes anonymous and underlying records confidential. Historical methods, including data swapping, cannot completely defend against the threats posed by today's technology. Enough noise must be added to protect confidentiality, but too much noise could damage the statistics' fitness-for-use. To help prevent anyone from tracing statistics back to a specific respondent, we will alter the underlying statistical tabulations before publication using a new method, called differential privacy, for the 2020 Census. This new, advanced, and far more powerful confidentiality protection system uses a rigorous mathematical process that protects respondents' information and identity by injecting random noise into the

tabulations. Differential privacy provides a way to control the tradeoff between noise and accuracy by creating a mathematical relationship between the noise added, the resulting privacy loss, and the resulting accuracy. It is also called "formal privacy" because it provides provable mathematical guarantees about the confidentiality protections that can be independently verified without compromising the underlying protections. Differential privacy is based on the cryptographic principle that an attacker should not be able to learn any more about an individual from the statistics the Census Bureau publishes using an individual's data than from statistics that did not use an individual's data.

On November 8, 2018, DSEP made several decisions related to the design of the 2018 E2E Census Test DAS and decided on a Privacy-Loss Budget, denoted by epsilon, for the test. These decisions are reflected in the prototype P.L. 94-171 file produced from the data collected during the test.

During its review, DSEP considered recommendations from the Chief Scientist and Associate Director for Research and Methodology, and reviewed many supporting materials that articulated the impact of invariants (as enumerated or exact statistics from the microdata) on differentially private systems, and illustrations of the accuracy privacy-loss tradeoffs of several values of epsilon.

DSEP made its decisions under the following assumptions:
- In deciding to implement differentially private methods, the Census Bureau has committed to trading privacy for accuracy only when there is a statutory mandate or clear public benefit to do so.
- There are no statutorily mandated uses of the prototype P.L. 94-171 file nor will these data be used for redistricting or enforcement of the Voting Rights Act.
- The primary uses of the prototype data are for testing systems and code, as well as informing stakeholders, and therefore the data do not require any significant degree of accuracy.
- Table consistency is a feature of the systems being designed for the 2020 Census and is not impacted by any decision regarding invariants.
- Since the test was restricted to Providence, RI, the highest level of geography that will be published is the county.

Based on those considerations and assumptions, DSEP made the following decisions with regards to the design of the 2018 E2E Census Test DAS.
1. The total population will be reported as enumerated (invariant) at the county level.
2. Disclosure avoidance will be applied to the voting age population at all levels of geography.
3. Number of housing units, number of occupied housing units, and number of group quarters facilities by group quarters facility type will be reported as enumerated (invariant) at the block level.
4. The Privacy-Loss Budget for the 2018 E2E Census Test will be an epsilon of 0.25. The Census Bureau will include a disclaimer for the data that notifies users that the data were produced to test systems and code, and the data may not be sufficiently accurate for any other uses.

All of these decisions apply exclusively to the 2018 E2E Census Test DAS and do not extend to the 2020 Census DAS. DSEP will decide on the design parameters for the 2020 Census DAS separately, and decide on a value of epsilon at a later date.

The 2020 Census Memorandum Series

The 2020 Census Memorandum Series documents significant decisions, actions, and accomplishments of the 2020 Census Program for the purpose of informing stakeholders, coordinating interdivisional efforts, and documenting important historical changes.

A memorandum generally will be added to this series for any decision or documentation that meets the following criteria:
   1. A major program-level decision that will affect the overall design or have significant effect on the 2020 Census operations or systems.
   2. A major policy decision or change that will affect the overall design or significantly impact 2020 Census operations or systems.
   3. A report that documents the research and testing for 2020 Census operations or systems.

Visit 2020 Census on Census.gov to access the Memorandum Series, the 2020 Census Operational Plan, and other information about preparations for the 2020 Census.